DOCUMENT RESUME

ED 448 422 CS 014 206

AUTHOR Diamond, Pollyann J.; Onwuegbuzie, Anthony J.
TITLE Short-Term Effects of Balanced Reading Implementation on Reading Achievement and Attitudes among Elementary School-Aged Students.
PUB DATE 2000-10-19
NOTE 46p.; Paper presented at the Annual Meeting of the Georgia Educational Research Association (Morrow, GA, October 19, 2000).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Black Students; Elementary Education; *Instructional Effectiveness; *Reading Achievement; *Reading Attitudes; *Reading Improvement; *Reading Instruction; White Students
IDENTIFIERS African Americans; *Balanced Reading Instruction; Georgia

ABSTRACT

A 4-Block Balanced Reading Program was implemented in a school system in Georgia during the 1998-99 school year in an effort to improve reading achievement of these students. This study examined the short-term effect of this program implementation on reading achievement and attitudes toward reading of students in Grades 1 through 5 (n=2,127) who had been exposed to this approach for 1 school year. Surprisingly, despite the fact that the 4-Block Balanced Reading Program has been utilized for nearly a decade, scant research exists on the efficacy of this reading intervention. Statistically significant reductions in achievement scores were found for Grade 5. Conversely, for Grade 4, a statistically significant increase was found for African American males. Cognitive Ability Test scores also were compared across grades. The fact that the fourth-grade scores did not differ from the earlier grades suggests that intelligence does not explain the relatively superior reading achievement of fourth graders over the 1-year period. With respect to reading attitudes, statistically significant differences were noted for Grade 2 (African American females), Grade 3 (African American females and males), Grade 4 (African American females and males), and Grade 5 (African American males and White females). In all cases, posttest scores were lower than were pretest scores, with moderate effect sizes, in general. Thus, the current findings, although preliminary, cast serious doubt over the effectiveness of the 4-Block Balanced Reading Program. Implications are discussed. Contains 28 references, 8 tables, and 3 figures of data. (Author/RS)

Reproductions supplied by EDRS are the best that can be made
from the original document.

TM

ED 448 422

Running head: READING ACHIEVEMENT

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

A. Onwuegbuzie

1

Short-Term Effects of Balanced Reading Implementation on Reading Achievement and Attitudes Among Elementary School-Aged Students

Pollyann J. Diamond

Anthony J. Onwuegbuzie

Valdosta State University

Correspondence should be addressed to: Anthony J. Onwuegbuzie, Department of Educational Leadership, College of Education, Valdosta State University, Valdosta, Georgia, 31698. E-Mail: tonwuegb@valdosta.edu

Paper presented at the annual meeting of the Georgia Educational Research Association, Morrow, Georgia, October 19, 2000.

CS 014 206

BEST COPY AVAILABLE

Abstract

A 4-Block Balanced Reading Program was implemented in a school system in Georgia during the 1998-99 school year in an effort to improve reading achievement of these students. This study examined the short-term effect of this program implementation on reading achievement and attitudes toward reading of students in Grades 1 through 5 (n = 2,127) who had been exposed to this approach for one school year. Surprisingly, despite the fact that the 4-Block Balanced Reading Program has been utilized for nearly a decade, scant research exists on the efficacy of this reading intervention. Statistically significant reductions in achievement scores were found for Grade 2 (African American and White females and males) and Grade 3 (African American females and males; White males); no significant differences were found for Grade 5. Conversely, for Grade 4, a statistically significant increase was found for African American males. Cognitive Ability Test scores also were compared across grades. The fact that the fourth-grade scores did not differ from the earlier grades suggests that intelligence does not explain the relatively superior reading achievement of fourth graders over the one-year period. With respect to reading attitudes, statistically significant differences were noted for Grade 2 (African American females), Grade 3 (African American females and males), Grade 4 (African American females and males), and Grade 5 (African American males and White females). In all cases, posttest scores were lower than were pretest scores, with moderate effect sizes, in general. Thus, the current findings, although preliminary, cast serious doubt over the effectiveness of the 4-Block Balanced Reading Program. Implications are discussed.

Short-Term Effects of Balanced Reading Implementation on Reading Achievement and Attitudes Among Elementary School-Aged Students

Reading achievement is of continuing interest and concern among educators (Snow, Burns, & Griffin, 1998). Much of the early research concerned the effectiveness of various materials and methods, but no consensus was reached as to the best single way to teach beginning readers (Bond & Dykstra, 1967, 1997). Interestingly, as early as 1967, Bond and Dykstra reported that no method of teaching reading was superior to any other, but that nonbasal programs tended to be superior to basal programs.

In a follow-up study, based on their investigation of the 27 projects that were part of the Cooperative Research Studies on First-Grade Reading Instruction, Bond and Dykstra (1997) reported that scores on some readiness subtests were positively related to reading achievement, but class size was determined to be unrelated to reading achievement. Teacher experience and efficiency ratings were statistically significantly but only slightly related to student achievement in reading. A combination of programs was found to be better than any one single approach, by systematic word skill instruction was essential regardless of the approach utilized. The relative success of non-basal programs compared with basal programs indicated that reading instruction might be improved by adopting elements from each of the approaches studied and the addition of a writing component was found to be effective. They further recommended that future research focus on teacher and learning situation characteristics rather than methods and materials.

In agreement with Bond and Dykstra (1967, 1997), Allington (1994) described the type of school that he considered would be needed for the 21st century. The importance of the classroom teacher and lessons in developing literacy, having long blocks of time to teach, and curricula that cover fewer topics with more depth using intensive instruction in literacy that integrate writing and reading were at the top of his list. Freppon and Dahl (1998) also supported a balance between skill instruction and whole language-like instruction in a constructivist classroom environment as the most effective approach to teaching reading.

Cassidy and Cassidy (1999) and Cassidy and Wenrich (1999) reviewed survey results from 1997 and 1998, respectively. In the original study, 25 members of the International Reading Association were chosen in proportionate numbers to membership worldwide and interviewed regarding the "hot topics" for literacy research; these individuals represented equal proportions of practitioners and university personnel. The second study surveyed the same group of educators. In both surveys, Balanced Reading was rated by 100% of the respondents as the top and second "hot topic" in the respective years.

Thus, considerable documentation exists to support the need for a balanced approach to teaching beginning readers (Allington, 1994; Cassidy & Cassidy, 1999; Cassidy & Wenrich, 1999; Cunningham & Allington, 1994; Freppon & Dahl, 1998). One such framework described in the literature is the 4-Block Balanced Reading framework (Cunningham, Hall, & Defee, 1991). Implemented initially in an elementary school in North Carolina, the framework divides the two-hour Language Arts instructional period

into four 30-minute blocks: Basal, Self-Selected Reading, Writing, and Working with Words. Results after the first year of implementation were promising, and, therefore, two additional sites, both in South Carolina, were added in subsequent years. What follows is a summary of the results at each of these three sites reported by Cunningham, Hall, and Defee (1998) since the implementation of 4-Block Balanced Reading.

School 1, a large suburban school in North Carolina, encompassed a student body with 20% to 25% low socioeconomic status (i.e., free/reduced price lunch) and 25% to 30% minority students. All classes were heterogeneous and each included approximate 23 students. No children were retained and none were referred for special education classes until second grade. The population was relatively stable, with approximately 10% turnover. All students in Grades 1 and 2 were administered the Basic Reading Inventory on an annual basis across the six years of implementation. There was no control group, but the authors reported consistent results across all groups of 100 to 140 children assessed annually. The most interesting results were among the students who did not read at grade level at the end of the Grade 1. Half or more of these students read at or above grade level by the end of Grade 2. Moreover, standardized test data for the system on the same groups of children, taken in Grades 3, 4, and 5, resulted in 90% of the students performing in the top two quartiles with respect to reading achievement (Cunningham et al., 1998).

School 2, also suburban but located in South Carolina, contained a student body with 25% low socioeconomic status. Approximately one-half of the teachers implemented the Four-Block framework. Again the Basic Reading Inventory was

administered. Students in classes where the Four Block Balanced Literacy Model was implemented scored at or above grade level at the end of the first semester of implementation; students in classes where the model was not implemented scored at grade level 1.2. Further, 100 of the students in classrooms where the model was implemented were matched with 100 students in classes where the model was not implemented. These students were matched on scores they received on the Cognitive Skills Assessment Battery, a readiness test administered to all students at the beginning of the school year. Encouragingly, the students who were in the Four-Block classrooms scored statistically significantly higher on the Metropolitan Achievement Test administered at the end of the first year of implementation (Cunningham et al., 1998).

School 3, a rural school located in South Carolina, was rated as one of the worst schools in the state based on achievement. The school included 84% low socioeconomic level students based on lunch status. As noted by Cunningham et al. (1998), various emphases had been implemented across the years, with no increase in achievement scores: in 1991-92, "skill and drill " basic skills instruction was introduced; in 1992-93, whole language was tried; in 1993-94, cooperative learning was implemented; and in 1994-95, Rita Dunn Learning Styles approach was incorporated. Only 20% of first graders and 9% of second graders in 1994-95 school year scored above the 50th percentile. During the 1995-96 school year, the Four Block Balanced Literacy framework was taught to teachers and mandated for use in the school. At the end of the first year, 30% of the first-grade students and 38% of the second-grade students scored at or above the 50th percentile. By the end of the second year of

implementation, 1996-97, 46% of the first-grade students and 40% of the second-grade students scored at or above the 50th percentile. These gains in achievement were attributed to the implementation of the Four Block Balanced Literacy Model (Cunningham et al., 1998). Cunningham (1999) reports that additional data currently are being gathered via a website from schools nationwide that are using this model.

In light of these research findings, the 4-Block Balanced Reading framework (Cunningham et al., 1998) was implemented in an additional school system in the state of Georgia during the 1998-99 school year in an effort to improve the reading achievement of these students. Unfortunately, at this school district, both reading achievement and reading attitudes had been found to decrease with increased years of schooling (Diamond & Onwuegbuzie, in press). Thus, the purpose of the present investigation was to examine the short-term effect of this program implementation on reading achievement and attitudes toward reading of students in Grades 1 through 5 who had been exposed to this approach for one school year.

Surprisingly, despite the fact that the 4-Block Balanced Reading Program has been used for nearly a decade, a paucity of research exists as to the efficacy of this reading intervention. Indeed, the three studies discussed above appear to be the only published research in this area. Yet, before nationwide implementation of the 4-Block Balanced Reading Program can be justified, it is imperative that there is extensive support for this technique. For the present inquiry, baseline data, obtained prior to the implementation of the 4-Block Balanced Reading Program, were compared with posttest data obtained one year later. Program effectiveness and student outcomes

were of primary interest in the current research. As such, it was expected that the current results would add to the scant body of published literature on the efficacy of 4-Block Balanced Reading initiative.

Methodology

Participants

Participants comprised 2,127 children who were enrolled in four Kindergarten through third-grade schools and two fourth- and fifth-grade schools in a small inner city school district in Georgia. That is, all regular education students who were enrolled in these six schools in Grades 1 through 5, who participated in the March, 1998 and 1999 administrations of the Iowa Tests of Basic Skills were included in comparisons of pre- and post-intervention measures of achievement. Those who completed the Elementary Reading Attitude Survey in September, 1998 and April 1999 were included in pre- and post-intervention comparisons of attitudes toward reading. However, students who received services as students with intellectual disabilities were excluded from the sample because they did not consistently participate in the general curriculum and did not typically take the Iowa Tests of Basic Skills. Because of the small number of students (approximately 30) in the "Other" category, which included Hispanic, Native American, and Asian, these data were not included in the analysis. Demographic data for the achievement sample are reported in Table 1; demographic data for the attitudes sample are reported in Table 2.

Insert Tables 1 – 2 about here

Instruments

The Iowa Tests of Basic Skills is administered annually to all Kindergarten through eighth-grade students at the school system where the study took place. This instrument is a standardized test of academic achievement developed and published by The Riverside Publishing Company (The Riverside Publishing Company, 1994). Areas measured are Reading, Language, and Mathematics. Items are presented in a multiple-choice format. The score of interest in this study was the Reading Total. Normal Curve Equivalent scores, which were designed for comparisons across grades and test levels, were the scores used in the present study. The State of Georgia has adopted this particular test for use in all Georgia schools for state-mandated testing and has provided evidence that the instrument generates valid scores for use in the Georgia public schools (Georgia Department of Education, 2000). The Riverside Publishing Company (1994) reported reliability coefficients for the Reading Total, as measured by KR-20, for Grades 1 through 4 which ranged from .89 to .93.

The Elementary Reading Attitude Survey was developed by McKenna and Kear (1990). This survey yields three scores: attitudes toward recreational reading (Items 1-10), attitudes toward academic reading (Items 11-20), and full scale (summation of 20 items). The items on this Likert-type scale are weighted from 1 to 4, with a score of 4 representing the most positive response, and 1 representing the least positive response. Students choose one of four Garfield line drawings which are representative in expression from most to least favorable toward a given reading activity. Using norms which were initially established on a population of 18,185 students in 38 states in

Grades 1-6, McKenna and Kear (1990) recommended that raw scores between 41 and 80 be indicative of increasingly positive attitudes toward reading, and scores 40 or less be considered increasingly negative attitudes toward reading.

McKenna and Kear (1990) reported reliability coefficients (i.e., Cronbach's alpha) of .80 or higher, except for the Recreational subscale in Grades 1 and 2 (i.e., .74 and .78, respectively). A factor analysis using the norm data supported the presence of two discrete scales reflecting different aspects of reading ability. Evidence of construct validity also was provided via positive correlations between the Recreational subscale and owning a library card and checking books out of the library, and a negative correlation between this subscale and time spent viewing television. Likewise, construct validity of the Academic subscale was evidenced through high scores on this subscale obtained by high ability readers; high ability readers were identified by teacher nomination in the norming process. This survey tool was chosen by the Language Arts Curriculum Director to measure attitudes toward reading.

As recommended by many researchers (e.g., Onwuegbuzie, 1999; Onwuegbuzie & Daniel, 1999, 2000; Thompson & Vacha-Haase, 2000; Wilkinson & the APA Task Force on Statistical Inference, 1999), reliability coefficients always should be reported for the data at hand. Unfortunately, no reliability information was available for the Elementary Reading Attitude Survey for the current sample. Thus, instead, as recommended by Vacha-Haase, Kogan, and Thompson (2000), for each grade, the standard deviation of scores from the inducted study (i.e., McKenna & Kear, 1990) was compared to the standard deviation of scores for the present sample. Deviation scores,

using the inducted study as the baseline, were as follows: Grade 1 (0.63); Grade 2 (0.71); Grade 3 (0.68); Grade 4 (1.12); and Grade 5 (0.04). Interestingly, these differences were all positive, suggesting that the current sample's reading scores were more variable than that of the inducted sample across all grade levels. However, with proportion deviations ranging from .01 to .10 (M = 0.06, SD = 0.03), it was concluded that the current sample was not too dissimilar from the inducted sample with respect to score variation on the Elementary Reading Attitude Survey. Indeed, using Magnusson's (1967) formula, which is based on the reliability of the inducted sample and the standard deviations of the inducted and present samples, the predicted reliability of the present sample's reading attitude scores was as follows: .88 for Grade 1, .89 for Grade 2, .89 for Grade 3, .91 for Grade 4, and .89 for Grade 5.

Procedure

The Iowa Tests of Basic Skills was administered in March 1998 and 1999. Scores were available through a data management program called Student Data Management System (1997). At the request of the Language Arts Curriculum Director, the Elementary Reading Attitude Survey was administered by classroom teachers during the week of September 21-25, 1998 and April 26-30, 1999 to all first- through fifth-grade students in the six schools. These data were collated, summed into the two subscales and a total score, and matched with Iowa Tests of Basic Skills data for each student. This data-gathering tool also included a place for demographic information regarding ethnicity and gender. These data also were coded with the attitudinal and achievement scores for each child. Lunch status was coded from information generated

by the School Nutrition Program database. Once all data were collated, identifying information was removed.

Results

Descriptive statistics regarding attitudes toward reading and reading achievement for this sample are displayed in Tables 3 and 4, respectively. A 2 (gender) x 2 (ethnicity) x 2 (socioeconomic status) x 4 (grade level) factorial analysis of variance (ANOVA) was conducted to examine reading achievement as a function of gender, ethnicity, socioeconomic status, and grade level. Findings revealed (a) a main effect for grade [F (4, 2645) = 21.91, $p < .001$; $\omega^2 = .18$]; (b) a main effect for ethnicity [F (1, 2645) = 154.58, $p < .001$; $\omega^2 = .24$]; and (c) a main effect for SES [F (1, 2645) = 98.67, $p < .001$; $\omega^2 = .19$]. No 2-, 3-, or 4-way interactions were noted. A post-hoc Scheffé analysis revealed that, students in Grade 1 outscored those in Grades 2 through 5, and students in Grade 2 outscored those in Grades 3 and 5. Also, White students outscored African American students. Finally, students paying full price for lunch outscored those who received free or reduced lunch.

Insert Tables 3– 4 about here

Using Cohen's (1988) criteria, the effect sizes, as measured by ω^2, were moderate for the ethnicity main effect and small for the grade and socioeconomic status main effects. Moreover, a cubic trend was found for grade [F (1, 2683) = 4.55, $p < .05$], with reading achievement declining consistently across Grades 2 and 3, rising slightly at

Grade 4, and decreasing slightly between Grades 4 and 5. The trend for reading achievement is shown in Figure 1.

Insert Figure 1 about here

A 2 (gender) x 2 (ethnicity) x 2 (socioeconomic status) x 5 (grade level) factorial analysis of variance (ANOVA) was conducted to examine attitudes toward reading as a function of gender, ethnicity, socioeconomic status, and grade level. Findings indicated (a) a grade x SES x ethnicity interaction [F (4, 2555) = 2.91, $p < .05$; $\omega^2 = .07$]; (b) a main effect for gender [F (1, 2555) = 40.54, $p < .001$; $\omega^2 = .13$]; (c) a main effect for ethnicity [F (1,2555) = 8.75, $p < .05$, $\omega^2 = .05$]; and (4) a main effect for grades [F (4, 2555) = 18.34, $p < .001$; $\omega^2 = .17$]. Using Cohen's (1988) criteria, the effect sizes, as measured by ω^2, were moderate for the main effect of ethnicity and small for all other main effects and for the interaction. Examination of the three-way interaction (grades x SES x ethnicity) plots revealed a steady monotonic decline for both White and African American females with White females consistently reporting more positive toward reading at each grade level. However, for males, the picture was different. Beginning in Grade 1, African American males were more positive than were White males, but by Grade 2, the situation had reversed. From that point, both groups steadily became less positive toward reading, with scores converging at Grades 3 and 5. The interaction pots are shown in Figure 2.

Insert Figure 2 about here

At all grade levels, females were more positive toward reading than were males and, at all levels except Grade 2, Whites were more positive toward reading than were African Americans. At Grade 1, African Americans were more positive toward reading than were Whites but only by less than one point, on the average. A post-hoc Scheffé analysis of the main effect for grade revealed that students in the first grade had statistically significantly more positive reading attitudes than did students in Grades 2 through 5; second- and third-grade students had statistically significantly more positive reading attitudes than did students in Grades 4 and 5; and fourth-grade students had statistically significantly more positive attitudes than did students in Grades 5. Moreover, a cubic trend emerged for grades [F (1, 2593) = 5.05, $p < .05$], with reading attitudes decreasing very slightly between Grades 1 and 2, decreasing sharply between Grades 2 and 3, slightly declining between Grades 3 and 4, and between Grades 4 and 5. The trend for reading attitudes is shown in Figure 3.

Insert Figure 3 about here

Comparisons were made between pre- and post-test scores in reading achievement using a series of dependent t-tests by grade, ethnicity, and gender. A Bonferroni adjustment was used to control for familywise error. Specifically, within each

grade, an alpha value of .0125 (i.e., .05/4) was utilized (Onwuegbuzie & Daniel, 2000). For Grade 1, achievement tests were not compared because pre-test Reading Total scores were not available. Table 5 presents descriptive statistics, including the results of dependent t-tests.

Insert Table 5 about here

For students in Grade 2, statistically significant differences were found for African American females, African American males, White females, and White males. All four groups scored significantly lower on the posttest than on the pretest: African American females (M = -10.5, SD = 14.19); African American males (M = -8.5, SD = 15.26); White females (M = -8.52, SD = 12.71); and White males (M = -7.79, SD = 15.64). Using Cohen's (1988) criteria, the effect sizes pertaining to these differences, which ranged from .42 to .67, were moderate for all four groups (Table 5).

For students in Grade 3, statistically significant differences were found for African American females, African American males, and White males. For all three groups, posttest scores were significantly lower than on the pretest scores: African American females (M = -5.75, SD = 14.34); African American males (M = -4.38, SD = 14.34); and White males (M = -7.22, SD = 12.77). Using Cohen's (1988) criteria, the effect sizes, ranging from .26 to .35, were small for all the groups (Table 5).

Interestingly, for Grade 4, all means were higher for the posttest than for the pretest, although the differences were not statistically significant, except for African

American males. Specifically, African American males scored significantly higher on the posttest (M = 3.0, SD = 13.75) than on the pretest. Using Cohen's (1988) criteria, the effect size was small for African American Males (Table 5). No significant differences were found for Grade 5 (Table 5).

To investigate why the 4-Block Balanced Reading implementation may have been more effective in Grade 4 in comparison to other grades, Cognitive Ability Test scores were compared across Grades 2 through 5. Fifth graders had higher Verbal Standard Age Scores [F (3, 1839) = 16.32, $p < .001$, $\omega^2 = .16$] than did each of the other four grades. However, no differences were noted among the Grades 2 through 4 (Table 6). Also, no differences were noted for the Composite Standard Age Scores (Table 7). The fact that the fourth-grade scores did not differ from the other three grades suggests that intelligence does not explain the relatively superior reading achievement of fourth graders over the one-year period.

Insert Tables 6-7 about here

Similarly, pre- and post-test scores on the Early Reading Attitudes Survey (McKenna & Kear, 1990) were analyzed using a series of dependent t-tests by grade, ethnicity, and gender. Again, the Bonferroni adjustment to control for familywise error was applied and an alpha level of .0125 (i.e., .05/4) was set. These results are presented in Table 8.

Insert Table 8 about here

Statistically significant differences were noted in Grade 2 for African American females; in Grade 3 for African American females and African American males; in Grade 4 for African American females and African American males; and in Grade 5 for African American males and White females. In all cases, posttest scores were lower than were pretest scores: Grade 2 African American females (M = -2.88, SD = 10.31); Grade 3 African American females (M = -3.08, SD = 10.34) and males (M = -3.02, SD = 11.94); Grade 4 African American females (M = -3.79, SD = 11.92) and males (M = -3.15, SD = 11.94), and Grade 5 African American males (M = -3.50, SD = 12.92) and White females (M = -4.56, SD = 9.42) (Table 6).

Discussion

The major purpose of the present study was to examine the short-term effects of the Four-Block Balanced Reading framework on reading achievement and attitude toward reading as a function of gender, ethnicity, socioeconomic status, and grade level among elementary school children. The use of a relatively large sample of students enrolled in a school system afforded the opportunity to add to the body of literature as related to that geographic area because of its high proportion of African American students and large number of students classified as low socioeconomic status. A further contribution of the current inquiry stems from the disaggregation of the data.

As in the baseline study (Diamond & Onwuegbuzie, in press), the trend for reading achievement and attitudes toward reading is for both to decrease with increased years of schooling. First-grade students appear to be the highest achievers and to have the most positive attitudes toward reading, with both reading indicators decreasing at each successive grade level. Interestingly, however, differences in intelligence did not seem to predict these lower reading scores among successively higher grade levels because no differences were noted among Grades 1 through 4 in Verbal IQ. Also, Grade 5 students, who had the highest average Verbal IQ scores, scored the lowest in both reading achievement and attitudes toward reading. Therefore, other factors should be examined to help clarify these trends.

The moderate effect sizes found in the current study, although preliminary, cast serious doubt over the effectiveness of the 4-Block Balanced Reading framework in improving both attitudes toward reading and reading achievement—at least in its first year of implementation. Cunningham et al. (1998) described hugely successful implementation projects in their studies completed in North and South Carolina. Students in these classrooms where the framework was implemented greatly increased reading achievement scores within the first year of its use. In the current study, pre- and post-test comparisons suggest that the converse is true for this population. Although population demographics are not identical between the Cunningham et al (1998) studies and the current population, this is unlikely be the only factor affecting outcomes resulting after implementation. To the extent that these findings are replicable, it is

possible that inadequate implementation of the 4-Block Balanced Reading framework has played role, at least in part, in preventing short-term gains from being realized.

Moreover, Onwuegbuzie (2000) conceptualized that implementation is a common and serious threat to internal validity in many educational intervention studies. According to Onwuegbuzie (2000), this threat often stems from differential selection of teachers. In particular, as the number of instructors involved in an instructional innovation increases, so does the likelihood that at least some of the teachers will not implement the initiative to its fullest extent. Such lack of adherence to protocol on the part of some teachers might stem from lack of motivation, time, or resources, inadequate knowledge, poor self-efficacy, implementation anxiety, stubbornness, poor attitudes, and the like (Onwuegbuzie, 2000). Thus, it is possible that such a lack of adherence to the 4-Block Balanced Reading protocol at the school district under investigation explains, at least in part, the reductions in reading achievement. Thus, future research should investigate whether implementation threat permeates the school district. One way of examining the level of implementation threat would be to compare the level of implementation of the 4-Block Balanced Reading Program between Grade 4 classrooms, where increases were observed, and the remaining grade levels, where significant decreases were apparent.

Attitudes toward reading also have continued to become more negative. Interestingly, this was true across all grades and for all demographically-defined groups. The interaction among grade, ethnicity, and gender also was interesting. Whereas the females, regardless of ethnicity, were steadily becoming less positive toward reading as

they advanced in grade, the picture for males was less consistent, with sharp, parallel decreases in attitudes noted between some grades for both ethnic groups and convergence at other grades. White females were overall more positive at every grade than were African American females, but White males were more positive than were African American males only at Grades 2 and 4. Scores for the two groups of males converged at Grades 3 and 5, with African American males reporting more positive attitudes toward reading than did White males at Grade 1. The reason for this difference is unknown and bears further investigation. In any case, the present findings highlight the utility of disaggregating data.

It is possible implementation threat also may have adversely affected the internal validity of the findings with respect to reading attitudes. For example, poor attitudes of some of the teachers toward the 4-Block Balanced Reading Program may have transgressed to their students. This should be the subject of future research.

Another component of the implementation threat that may have prevailed is related to time. That is, it is possible that one year may not be a sufficient period of time to observe positive gains in reading achievement and attitudes. However, significant increases in reading performance have been observed by the end of the first year of introduction of the 4-Block Balanced Reading Program (Cunningham et al., 1998). Moreover, what is particularly disturbing is that reading achievement decreased for four of the five grades. No significant differences between the pre- and post-intervention measures perhaps could have been overlooked; however, decrements in reading outcomes must be taken seriously, thereby justifying much more research in this area.

A framework that offers great potential for studying the role of implementation threat in the 4-Block Balanced Reading Program is that pertaining to Rogers (1995). In his seminal work, entitled, The Diffusion of Innovations, Rogers (1995) identified the following five stages individuals go through in deciding whether to adopt a new innovation: (a) knowledge (e.g., awareness and comprehension), (b) persuasion (e.g., formation of attitude), (c) decision (e.g., trial, followed by adoption or rejection), (d) implementation (i.e., actual use), and (e) confirmation (i.e., continued use). Thus, researchers in the future should consider comparing teachers whose students experience gains in reading achievement after implementation of the 4-Block Balanced Reading Program to their lower-achieving counterparts with respect to these five stages. These five stages also could be analyzed via narrative profile formation techniques (Tashakkori & Teddlie, 1998; Witcher, Onwuegbuzie, & Minor, in press). For example, average narrative profiles could be compared for these two sets of teachers (Witcher et al., in press).

If differences are found in the degree and speed that teachers adopt the Balanced Reading Program, then the characteristics of the organization (i.e., the local school district under study) (Rogers, 1995) should be examined further. In addition, as recommended by Rogers, the attributes of the innovation should be studied. Rogers (1995) identified the following five major attributes that influence how quickly an innovation is likely to be adopted: (a) relative advantage, (b) compatibility, (c) complexity, (d) trialability, and (e) observability. According to this author, innovations are likely to be adopted more slowly if (a) they offer advantages such as high initial cost,

relative economic disadvantages, increase in discomfort, savings in time and effort, and social status; (b) if they are not compatible with the values, beliefs, perceived needs, and goals of the teachers and administrators, as well as the parents and other stakeholders; (c) if they are difficult to understand, to implement, and to utilize; (d) if they are not phased in gradually; and (e) if they do not produce invisible results. Thus, each of these components should be scrutinized.

In summary, future research should focus on issues affecting implementation of the 4-Block Balanced Reading framework. As noted above, demographic differences between the various samples reported in the literature and the current sample is unlikely to be the sole reason for different outcomes. Moreover, these results suggest further questions that would be better investigated using qualitative methods. These questions include the following: To what extent is the 4-Block Balanced Reading Program being implemented? What has been implemented differently in fourth-grade classes as compared with the other grades? What needs to be changed to reverse the short-term downward trend observed in reading achievement so that these students can compete in the information society in which they must function as adults? Clearly, answers to such questions are vital and must be investigated. Questions such as these must be asked in a society where accountability for achievement and expenditure of funds is such a significant issue.

References

Allington, R. (1994). The schools we have. The schools we need. The Reading Teacher, 48, 14-29.

Bond, G., & Dykstra, R. (1967). The cooperative research program in first-grade reading instruction. Reading Research Quarterly, 2, 5-142.

Bond, G., & Dykstra, R. (1997). The cooperative research program in first-grade reading instruction. Reading Research Quarterly, 32, 348-427.

Cassidy, J., & Cassidy, D. (1999). What's hot, what's not for 2000. Reading Today, 17, 1-2.

Cassidy, J., & Wenrich, J. (1999). Literacy research and practice: What's hot, what's not, and why. Reading Teacher, 54, 402-406.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). New York: John Wiley.

Cunningham, P. (1999, July 27). Research. [WWW document] URL http://wfu.eduction/~cunningh/fourblocks/Research.htm.

Cunningham, P., & Allington, R. (1994). Classrooms that work: They can all read and write. New York: Harper Collins.

Cunningham, P., Hall, D., & Defee, M. (1991). Nonability-grouped multilevel instruction: A year in a first grade classroom. The Reading Teacher, 44, 566-571.

Cunningham, P., Hall, D., & Defee, M. (1998). Nonability-grouped multilevel instruction: Eight years later. The Reading Teacher, 51, 652-664.

Diamond, P., and Onwuegbuzie, A. (in press). Factors associated with reading achievement and attitudes among elementary school-aged students. Research in the Schools.

Freppon, P., & Dahl, K. (1998). Balanced instruction: Insights and considerations. Reading Research Quarterly, 33, 240-251.

Georgia Department of Education (2000, July 9). *1998-99 Georgia public education report card*. [WWW document] URL http://accountability.doc.k12.ga.us/Report99/default.html.

Magnusson, D. (1967). Test theory. Boston, MA: Addison-Wesley.

McKenna, M., & Kear, D. (1990). Measuring attitude toward reading: A new tool for teachers. The Reading Teacher, 43, 626-639.

Onwuegbuzie, A.J. (1999, September). Common analytical and interpretational errors in educational research. Paper presented at the annual meeting of the European Conference on Educational Research, Lahti, Finland.

Onwuegbuzie, A.J. (2000, November). Expanding the Framework of internal and external validity in quantitative research. Paper presented at the annual meeting of the Association for the Advancement of Educational Research (AAER), Ponte Vedra, Florida.

Onwuegbuzie, A.J., & Daniel, L.G. (1999, November). Uses and misuses of the correlation coefficient. Paper presented at the annual conference of the Mid-South Educational Research Association, Point Clear, AL.

Onwuegbuzie, A.J., & Daniel, L.G. (2000, April). Common analytical and interpretational errors in educational research. Paper presented at the annual conference of the American Educational Research Association, New Orleans.

Rogers, E.M. (1995). The diffusion of innovations (4th ed.). New York: The Free Press.

Snow, C., Burns, S., & Griffin, P (Eds.). (1998). Preventing reading difficulties in young children. Washington, D.C.: National Academy Press.

Student Data Management System [Computer software]. (1997). Chicago, IL: The Riverside Publishing Company.

Tashakkori, A., & Teddlie, C. (1998). Mixed methodology: Combining qualitative and quantitative approaches. Applied Social Research Methods Series (Vol. 46). Thousand Oaks, CA: Sage.

The Riverside Publishing Company (1994). Technical summary I: Riverside 2000. Chicago, IL: The Riverside Publishing Company.

Thompson, B., & Vacha-Haase, T. (2000). Psychometrics is datametrics: The test is not reliable. Educational and Psychological Measurement, 60, 174-195.

Vacha-Haase, T., Kogan, L.R., & Thompson, B. (2000, April). Sample compositions and variabilities in published studies versus those in test manuals: Validity of score reliability inductions. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Wilkinson, L., & the Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. American Psychologist, 54, 594-604.

Witcher, A., Onwuegbuzie, A.J., & Minor, L. (in press). Characteristics of effective teachers: Perceptions of preservice teachers. Research in the Schools.

Table 1

Demographic Distribution of Achievement Sample as a Function of Grade Level (Grades 2-5)

Grade	African American	White	Male	Female	Total
2	77.1%	22.9%	48.2%	51.8%	
	357	106	223	240	463
3	78.2%	21.8%	45.6%	54.4%	
	348	97	203	242	445
4	77.6%	22.4%	49%	51%	
	326	94	206	214	420
5	76.8%	23.2%	50.8%	49.2%	
	301	91	199	193	392
Total	77.4%	22.6%	7/.3%	51.7%	
	1332	388	831	889	1720

Table 2

Demographic Distribution of Attitudes Sample as a Function of Grade Level (Grades 1-5)

Grade	African American	White	Male	Female	Total
1	74.8%	25.2%	51.8%	48.2%	
	357	120	247	230	477
2	77.8%	22.2%	49.4%	50.6%	
	364	104	231	237	468
3	76.7%	23.3%	47.6%	52.4%	
	348	106	216	238	454
4	75.7%	24.3%	49.3%	50.7%	
	281	90	183	188	371
5	73.7%	26.3%	52.5%	47.5%	
	261	93	186	168	354
Total	75.8%	24.2%	50.04%	49.96%	
	1611	513	1063	1061	2124

Table 3

Spring 1999 Total Reading Means and Standard Deviations by Grade, Gender, and Ethnicity

	Male				Female				
	African American		White		African American		White		
Grade	M	SD	M	SD	M	SD	M	SD	n
1	50.79	17.44	64.72	19.49	51.57	18.28	71.30	18.20	586
2	42.79	16.66	63.51	20.62	46.34	14.97	60.70	16.42	585
3	38.36	15.98	62.79	19.91	40.68	15.78	61.85	17.36	539
4	40.85	17.11	56.59	18.35	40.85	15.87	64.21	17.57	499
5	38.81	15.53	58.32	16.03	38.63	15.27	58.53	14.00	476
Total	42.67	17.21	61.44	19.05	43.84	16.78	63.45	17.33	2685

Table 4

Spring 1999 Total Attitude Means and Standard Deviations by Grade, Gender, and Ethnicity

	Male				Female				
	African American		White		African American		White		
Grade	M	SD	M	SD	M	SD	M	SD	n
1	63.20	12.24	61.11	14.68	66.5	10.68	67.64	10.27	582
2	60.86	11.33	63.00	10.11	65.67	9.95	67.17	9.95	551
3	58.65	12.30	58.82	10.72	63.79	10.56	64.90	8.40	533
4	54.26	12.55	58.04	11.00	59.20	10.83	61.95	9.57	484
5	53.19	11.01	53.69	9.81	58.69	11.84	60.64	8.58	445
Total	58.41	12.48	58.88	11.96	63.09	11.33	64.76	9.78	2595

Table 5

Spring 1999 and Spring 1998 Total Reading Means, Standard Deviations, and t-Values by Grade, Gender, and Ethnicity

			Posttest (Spring 1999)		Pretest (Spring 1998)			Effect Size
Grade	Gender	Ethnicity	M	SD	M	SD	t	Cohen's d
2	Female	African American	46.65	14.70	57.15	16348	-9.81**	.67
2	Female	White	62.72	16.87	71.23	19.13	-5.36**	.47
2	Male	African American	42.41	16.75	50.99	18.44	-7.57**	.49
2	Male	White	64.57	20.19	72.36	16.54	-3.23*	.42
3	Female	African American	40.36	15.89	46.11	17.26	-5.56**	.35
3	Female	White	63.77	17.94	67.10	20.06	-1.79	
3	Male	African American	38.18	15.94	42.55	17.40	-3.70**	.26
3	Male	White	60.88	20.67	68.10	20.58	-3.96**	.35
4	Female	African American	41.00	16.02	39.31	12.71	1.96	
4	Female	White	63.35	17.01	59.67	17.54	2.26	

Note: * $p < .01$; ** $p < .001$

Table 5 (continued)

Spring 1999 and Spring 1998 Total Reading Means, Standard Deviations, and t-Values by Grade, Gender, and Ethnicity

			Posttest (Spring 1999)		Pretest (Spring 1998)			
Grade	Gender	Ethnicity	M	SD	M	SD	t	Effect Size (Cohen's d)
4	Male	African American	40.74	17.16	37.73	14.31	2.82*	.19
4	Male	White	59.10	18.92	54.75	17.34	2.21	
5	Female	African American	39.10	14.54	40.42	16.19	-1.30	
5	Female	White	58.49	12.08	61354	15.63	-1.81	
5	Male	African American	39.08	15.25	39.20	18.13	-.11	
5	Male	White	58.52	16.03	59.75	19.21	-.85	

Note: * $p < .01$; ** $p < .001$

Table 6

Cognitive Abilities Test: Verbal SAS Means and Standard Deviations by Grade, Gender, and Ethnicity

	Male				Female				
	African American		White		African American		White		
Grade	M	SD	M	SD	M	SD	M	SD	n
2	82.88	13.54	102.06	18.40	83.37	12.52	100.01	14.56	579
3	83.30	12.10	105.13	13.74	84.00	12.69	106.34	17.44	462
4	83.83	13.32	99.97	16.11	83.19	12.40	104.28	14.40	350
5	89.95	13.66	107.20	16.80	90.07	12.98	107.91	12.52	464
Total	84.94	13.52	104.06	16.52	85.15	12.94	103.93	15.05	1855

Table 7

Cognitive Abilities Test: Composite SAS Means and Standard Deviations by Grade, Gender, and Ethnicity

	Male				Female				
	African American		White		African American		Whites		
Grade	M	SD	M	SD	M	SD	M	SD	n
2	90.61	14.44	108.20	16.62	91.78	12.76	106.63	15.75	575
3	90.60	12.82	115.10	15.17	92.65	12.00	114.47	16.87	460
4	91.74	11.82	106.16	15.38	92.51	12.68	112.41	14.65	350
5	92.26	13.28	110.36	14.88	92.50	12.98	111.30	10.38	441
Total	91.22	13.33	110.23	15.74	92.34	12.56	110.52	15.09	1826

38

Table 8

Spring 1999 and Fall 1998 Total Attitude Means, Standard Deviations, and t-Values by Grade, Gender, and Ethnicity

			Posttest (Spring 1999)		Pretest (Spring 1998)			
Grade	Gender	Ethnicity	M	SD	M	SD	t	Effect Size (Cohen's d)
1	Female	African American	66.07	10.92	66.88	11.46	-.81	.
1	Female	White	67.06	10.77	66.91	11.46	.09	
1	Male	African American	62.388	12.29	62.90	12.10	-.01	
1	Male	White	61.42	13.57	63.82	13.05	-1.45	
2	Female	African American	65.38	10.84	68.70	10.29	-3.73**	.27
2	Female	White	68.53	7.91	69.95	9.03	-1.17	
2	Male	African American	60.56	11.52	60.54	13.25	.02	
2	Male	White	63.69	9.88	63.76	12.42	-.04	
3	Female	African American	63.92	10.54	66.99	10.47	-4.05**	.29
3	Female	White	64.72	8.63	65.68	11.37	-.74	

Note: * $p < .01$; ** $p < .001$

Table 8 (continued)

Spring 1999 and Fall 1998 Total Attitude Means, Standard Deviations, and t-Values by Grade, Gender, and Ethnicity

			Posttest (Spring 1999)		Pretest (Spring 1998)			
Grade	Gender	Ethnicity	M	SD	M	SD	t	Effect Size (Cohen's d)
3	Male	African American	58.73	12.34	61.75	11.47	-3.43*	.25
3	Male	White	59.38	10.85	62.06	12.17	-1.85	
4	Female	African American	58.63	11.05	62.43	12.17	-3.68**	.33
4	Female	White	62.07	9.69	64.17	9.66	-2.19	
4	Male	African American	53.61	12.27	56.76	11.85	-2.80*	.26
4	Male	White	58.22	10.65	55.53	13.28	-1.55	
5	Female	African American	59.47	11.94	61.31	10.64	-1.83	
5	Female	White	60.44	9.00	65.00	7.25	-2.90*	.56
5	Male	African American	53.20	10.52	56.71	11.62	-3.26*	.32
5	Male	White	53.98	10.08	57346	10.58	-2.18	

Note: * $p < .01$; ** $p < .001$

Figure 1

Estimated Marginal Means of Spring Reading Total

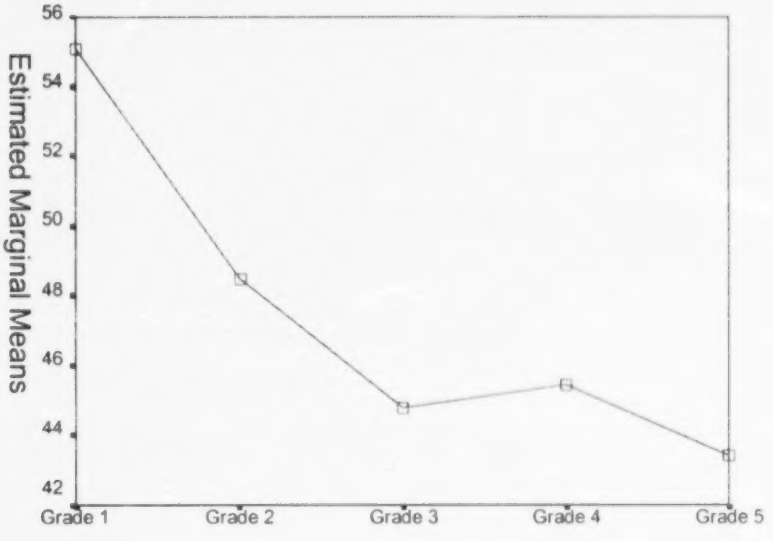
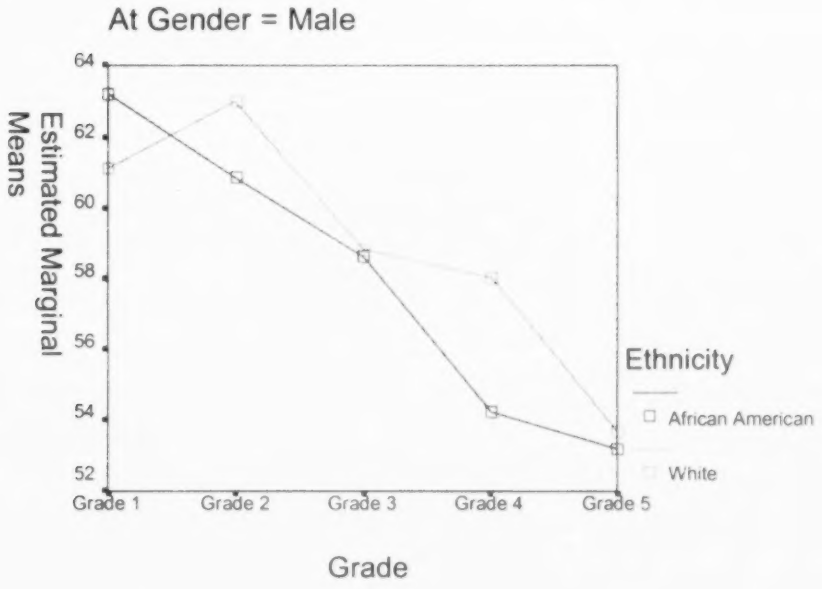


Figure 2

Estimated Marginal Means of Spring Total Attitude



Estimated Marginal Means of Spring Total Attitude

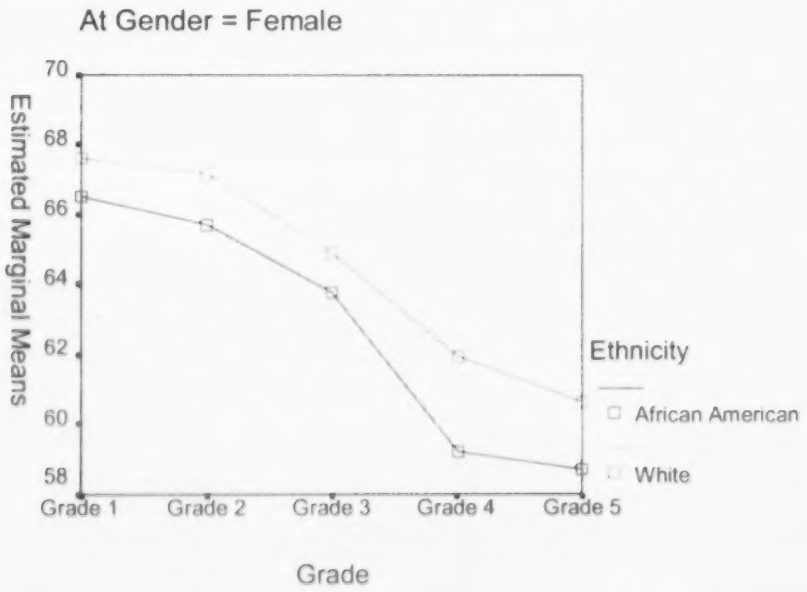


Figure 3

Estimated Marginal Means of Spring Total Attitude